Displays 39 (2015) 33–41

Contents lists available at ScienceDirect

Displays

journal homepage: www.elsevier.com/locate/displa

Identifying regions of interest in reading an image

Cheng-Min Tsai [a,b,*], Shing-Sheng Guan [a]

[a] Graduate School of Design, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan, ROC
[b] Department of Creative Product Design and Management, Far East University, Tainan 74448, Taiwan, ROC

ARTICLE INFO

Article history:
Received 16 August 2014
Received in revised form 19 July 2015
Accepted 2 August 2015
Available online 3 August 2015

Keywords:
Region of interest
Image quality assessment
Eye movement
Eye-tracking device

ABSTRACT

The aim of this study is to develop an effective method to analyze regions of interest (ROIs). Two experiments were conducted at different times using different groups of observers with different images on different displays. Observers' eye-movement data were collected. Fixation maps showing CIELAB L^* values were created. The ΔL^* values between the two maps were used to quantify differences in visual fields, counting methods, observer variability and repeatability between the two experiments.

The results showed that fixation maps can be used to effectively analyze the distribution of eye movements between images. The ΔL^* value calculated for two fixation maps is easy to understand and computes differences based only on ROIs more effectively than differences based on the entire image. The results from the two experiments were consistent, indicating that eye-tracking data are robust for evaluating image quality.

1. Introduction

One of the goals of imaging research is to develop a metric based on image statistics to assess images. Earlier metrics, such as the peak signal-to-noise ratio (PSNR) [1] and root mean square error (RMSE) [2], were focused on the physical measurement of image quality and do not correspond well with the results of visual assessment [3–5]. More recent studies were conducted based on visual assessments using psychophysical methods to investigate parameters that affect the judgment of image quality, such as the naturalness, colorfulness, and sharpness of the images. Efforts were then made to develop color models that predict these parameters and pool the individual parameters to form an overall image quality index [6–13].

The above methods were based on a global analysis of the entire image. Another approach is to identify the regions of interest (ROIs) in an image, defined as the areas of an image that attract more visual attention than the other areas [14,15]. Privitera and Stark have noted that ROIs are defined as the loci of the human's eye fixations and that they can be analyzed using their spatial distribution over the visual stimulus and their temporal ordering [14]. Privitera and Stark have developed an image quality model based on the focus of visual attention within an image rather than the entire image [15,16]. Privitera et al. [15] have conducted a series of experiments to fit the measured scan path data and identify ROIs, first using an eye-tracking system. They then used geometrical spatial kernels and linear filter models to locate the ROIs in an image. Nguyen et al. [3] have grouped ROIs based on an analysis of scan-paths and sequences of fixation for viewing grayscale images and subsequently performed compression based on the ROIs of an image. The algorithm only addresses grayscale images and thus may not work well for color images.

Because the eye is the first element of the visual system to receive visual information, it is also the only means by which the brain obtains external images. Henderson and Hollingworth [17] have found that eye movements are critical for the efficient and timely acquisition of visual information during complex visual-cognitive tasks. The eye-tracking technique has been widely applied in various research areas, such as human factor and interface design, advertising and marketing, psychology and neuroscience, attention span studies and visual text analysis. In the image assessment research field, fixation-map analysis provides an opportunity to objectively define the principal ROIs for viewing images [18]. This study differs from the previous studies in that it employs an eye-tracking technique to develop an effective method to analyze ROIs.

2. Experimental design

Two experiments for assessing image quality were conducted to test the reliability of the results from the eye-tracking experiment. These experiments were conducted at different times using

* Corresponding author at: Graduate School of Design, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan, ROC.
E-mail address: ansel.tsai@gmail.com (C.-M. Tsai).

http://dx.doi.org/10.1016/j.displa.2015.08.001
0141-9382/

different groups of observers with different images on different displays that have different sizes and peak white values. The results from the two experiments can be compared to test the robustness of the method. Table 1 summarizes the experimental conditions in these two experiments.

2.1. Eye-tracking technique

A device used for measuring eye movements is referred to as either an eye-tracking device or an eye tracker [19]. An EyeLink II (SR Research Ltd., Mississauga, Ontario, Canada) with an infra-red head-mounted, video-based, pupil and corneal reflection eye-tracking apparatus was used in this study. This device has high resolution (noise-limited at <0.01°) and a fast data-acquisition rate (more than 250 samples per second). It was easy to operate and required less than a minute to conduct the calibration process before commencing the experiment. Fig. 1 shows the experimental conditions. Observers were seated in front of an LCD-TV and wore a headset containing a camera to record their eye movements and fixation locations (see Fig. 2). The time that users fixated on each pixel and the eye positions for each image were stored on the Host PC (personal computer). The observers judged the image quality on a stimulus monitor controlled by the Display PC. The eye-movement data were processed instantly to yield the fixation position and fixation duration in a form ready for use in the data analysis.

2.2. Experiment I

2.2.1. Display

A 30-inch LCD TV was used to display images. A Gretag Macbeth Eye-One colorimeter was used to establish the ICC profile for display characterization. The study was conducted in a laboratory with an ambient illuminance of approximately 230 lux and a correlated color temperature (CCT) approaching 6500 K.

2.2.2. Observers

Thirty observers participated in Experiment I (19 females and 11 males whose average age was 23). All were staff members and postgraduate students from the School of Design at the National Yunlin University of Science and Technology. All had normal color vision according to the Ishihara color vision test.

Table 1
Conditions for Experiments I and II.

Materials	Experiment I	Experiment II
Observers	30	15
Images	11	6
Display	Sharp 30″ LCD-TV	SONY 40″ LCD-TV
Peak white (x, y chromaticity coordinates)	(0.3127, 0.3285)	(0.3153, 0.3297)
Peak white (luminance (cd/m^2))	140	203
Mid-gray background color	L^* = 61.96, a^* = -1.21, b^* = −13.21	L^* = 74.75, a^* = 1.45, b^* = −12.84
Resolution of the display	1280 × 768 pixels	1920 × 1080 pixels
Image size (pixels)	1280 (W) × 768 (H)	768 (W) × 512 (H)
Viewing distance	120 cm	70 cm
Viewing angle (degree)	≒30.56 (W) × 18.38 (H)	≒29.30 (W) × 19.21 (H)
The psychometric attributes scaled	Total image quality	Total image quality Brightness Saturation Naturalness Preference
Averaged ambient lighting	233 lux, ≒6500 K	233 lux, ≒6500 K

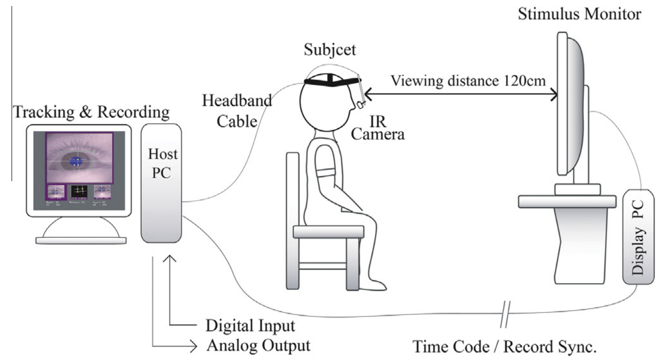

Fig. 1. The eye tracking apparatus set for experimental condition.

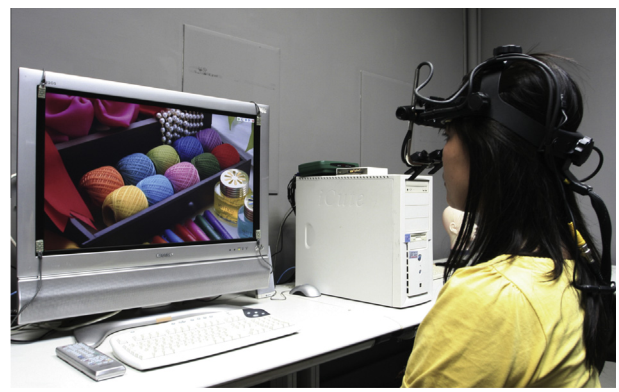

Fig. 2. The real condition of eye tracking experiment.

2.2.3. Experimental Images

The experimental images were selected from ISO standards (ISO 12640-1 (1997), ISO 12640-2(1997), ISO 12640-3(2004), ISO 12640-3(2007)), and some were collected from the *Kodak Lossless True Color Image Suite* [21]. All of these images were first categorized into three groups: portraits, landscapes with architectural images, and indoor multiple-object images. The experimental images were then selected from each category for the current experiment, as shown in Fig. 3. The size of an image was 1280 × 768 (pixels). The images were randomly displayed during each observation session against a background of a mid-gray color with an L^* of approximately 60.

2.3. Experiment II

Experiment II was performed to test the repeatability of the eye-tracking results. The environment for Experiment II was the same as that of Experiment I.

2.3.1. Display

A different display 40 in. in size (73.26° and 40.52° in the horizontal and vertical directions of full frame) was used. The viewing distance on the single image was adjusted to have the same visual field (29.30° and 19.21° in the horizontal and vertical directions, respectively).

2.3.2. Observers and experimental images

Fifteen observers (8 males and 7 females, whose average age was 27) participated in the experiment. Two of the subjects also took part in Experiment I. Fig. 4 shows the six images used, three of which were used in Experiment I.